CORE
Metadata, citation and similar papers at core.ac.uk
Provided by GSSRR.ORG: International Journals: Publishing Research Papers in all Fields

International Journal of Sciences: Basic and Applied Research (IJSBAR)

ISSN 2307-4531
(Print & Online)

http://gssrr.org/index.php?journal=JournalOfBasicAndApplied

Assessing Instrument Reliability and Validity in Exploring Teacher's Roles and Attitude in ESP

Arta Vinca*

SEEU, Tetovo 1200, Macedonia
Email: arta_vinca@yahoo.com

Abstract

This study aimed to validate a questionnaire on Exploring Teacher's Roles and Attitude in ESP Classes - a Case Study in Higher Education in Kosovo. Participants (n=215) were undergraduate students who have taken a Legal English course at the University of Prishtina for two semesters. Test reliability, internal reliability, and construct validity of the questionnaire were determined. Overall internal reliability (Cronbach's alpha= Cronbach's alpha = 0.8377) indicated a high level of internal consistency. Construct validity for the questionnaire was satisfactory as the analysis proved p-value of 0.08, indicating correlation at 90% level of confidence.

Keywords: ESP; Questionnaire; Reliability; Validity; Kosovo.

1. Introduction

A pilot study was conducted to test the reliability and validity of the questionnaire. The questionnaire was designed exclusively for the PhD study on "Exploring teacher attitudes and roles in ESP courses" in English. Teachers play a significant role via their attitude and the roles they must undertake under such settings. In this aspect there is not much reliable data that can be retrieved for regional and up-to date studies. Also most of the studies that focus on ESP they don't measure the importance of the teacher attitudes and their roles in an ESP Kosovar setting. Students need awareness about ESP courses and what roles teachers have.

---

* Corresponding author.

They need to know about 'needs analysis', ESP methodology, and ESP in general. Assessing and identifying the needs and the specific nature of the ESP course would help students learn more and prepare for the course on time. This study aids to the process of validation and the testing of reliability of a tool that asses the knowledge of the students about ESP and teacher attitudes and roles in it for Kosovar UP students. Given that the questionnaire proves good validity and reliability it can be used for the main study in assessing teachers' attitudes and their roles in an ESP course. A starting point was to identify the population of all those students that make up the University of Prishtina (UP). The school (faculty) of Law was identified as a representative school for the pilot study to be conducted. Therefore 215 students participated in questionnaire survey. The respondents were randomly selected using simple random sampling.

1.1. Reliability and Validity of the questionnaires

Reliability refers to the repeatability of findings. If the study were to be done a second time, would it yield the same results? If so, the data are reliable. If more than one person is observing behavior or some event, all observers should agree on what is being recorded in order to claim that the data are reliable.

All 41 Likert scale questions were tested for reliability. It is worth noting that out 215 observations (surveys) a total 139 were used as 73 had missing values. Nevertheless, results of the item analysis resulted in a Cronbach's alpha = 0.8377 indicating a high level of internal consistency.

Validity indicates the degree to which an instrument measures what is supposed to measure. Likert Scales are the most commonly used type of survey questions, yet one of the main weaknesses is that the validation may be difficult to demonstrate. Acknowledgement or consideration of difficulty in obtaining exact expectation due to variations between groups is a very important factor.

Content validity measures the extent a measure represents all aspects of a given social construct, such as the case in this study. This part of validity was measured through the study's lead expert who ensured that data were indeed measuring all sides of the research.

Convergent validity is defined as the extent to which the scores on one measure are related to scores collected from a similar or different measure.

The next step used in testing for convergent validity was to use the Spearman Rank correlation coefficient (rho) between the expected responses (as predicted by experts) and actual responses. Given that data gathered through Likert Scales questions are of ordinal type, the most frequent response (mode) was used in calculating rho and its p-value. Analysis resulted in a p-value of 0.08, indicating that these two groups are correlated at 90% level of confidence.

2. Rationale and objectives

The theoretical ground for the study is based on the most popular recent research on Second Language research: [1, 2, 3, 4, 5, 6]. According to Dörnyei [7] questionnaires are considered "any written instruments that present

respondents with a series of questions or statements to which they react either by writing out their answers or selecting from among existing answers." According to Mackey [8] questionnaires are structured data collection instruments:

"The typical questionnaire is a highly structured data collection instrument, with most data asking about very specific pieces of information or giving various response options for the respondents to choose from, for example by ticking a box or circling the most appropriate option."

Most professional questionnaires are made of "close ended" items, which do not require the respondents to produce any free writing. The most famous type of close-ended item is undoubtedly the Likert Scale, which consists of a characteristic statement accompanied by five or six response options for respondents to indicate the extent to which they "agree" or "disagree" with it by marking or circling one of the responses ranging from "strongly agree" to "strongly disagree".

For the questionnaire to prove validity and reliability it requires good sets of items in order to test hypothesis and give proper data analysis. For this matter good questionnaire items were created in a close correlation to the following terms: "short and simple" items, each consisting of maximum 20 words; the language that was used was strictly "simple and natural', "double-barreled questions were avoided; In contrary to [8] idea that "negative construction" should be avoid as it can be problematic for the respondents in the current questionnaire item containing negative construction (6), were developed. Due to the nature and the specific of the topic and the context where the piloting was conducted it was recommended to use negative construction in order to cross check items for validity and reliability purposes.

The design of the questionnaire plays a very significant role in responsiveness therefore it must be closely considered. Few important points, as literature review suggests are: length, space and economy, mixed scale and items, and factual (personal) items at the end [8].The questionnaire was constructed to fit four pages, because anything over 4-6 pages requires over 30 minutes to respond, which is considered as an "imposition" [8]. The fact that we want to keep the questionnaire to maximum 4 pages few modifications were made to the page setting: therefore margins were reduced, smaller font was used, and also the page was used to its whole width with responses reflecting next to the items. According to [8] the "personal" items should be at the end, however for the present study the personal information was presented at the first part as it was very important due to the test-re-test of the respondent at a later stage. Having in mind that the respondents may lose focus towards the end and not provide the crucial information would cause significant loses in the data analysis stage and later on in a further study of the same sample. According to [4] in order to meet the up-to date standards for a reliable L2 questionnaire the criteria needs to be set. First - the concept about the questionnaires, what they are, what can they measure, and their 'pro et contra' features were reviewed; Second - constructing the questionnaire.

3. Methodology

The ESP (English I and II) course at the Law faculty aims to cover the most important parts of legal English, as a foreign language. It covers Legal English concepts and terminology, as well as various language skills such

as: writing, reading, speaking, listening, presentation skills (through joint projects, Mock trials) etc. Upon completion students gain competence in understanding and become familiarized with real life situations that occur in the field of Law in real professional settings. The vocabulary is the backbone of every ESP course, but besides it a lot of attention is also payed and in associating the terminology and the legal concept with Albanian, where sometimes for better understanding translation and syntax comes handy also. The course introduces the Legal English only so that the students get the basic idea about the Law, the course mainly aims to teach them the skills that they will need in their further education and their future profession.

In a Kosovar context English is the major foreign language and the most widely spoken language. Due to its popularity and usage throughout Kosovo more specialized courses, such as ESP, also started to gain momentum at a tertiary level and vocational education, as well as in professional settings. Students from most academic units within the University of Prishtina register for either elective or compulsory English courses within their departments. Although they attend the courses their overall proficiency lags behind greatly even upon completion of the course. Due to this students from UP are obliged to take private English courses that will prepare them for their future studies or employment. There are different factors that impede students' proficiency and overall results which then reflect their fluency. To name just a few: poor language skills in the ESP field, teaching methodology and approach, teachers' attitude and their roles, over populated classes, lack of materials, etc. A closer research in this field will definitely ease current difficulties that the students are experiencing.

3.1. Participants

Participants in this study were first – year undergraduate students at University of Prishtina Law Faculty, Kosovo. The sample of this study included the students that had taken English I and English II (English for Lawyers) course at the Law Faculty during the academic year 2014-2105 (n=215). All of the participants gave verbal consent to participate in the study and to fill up the questionnaire. UP is a public University in Kosovo that besides Prishtina it is stretched out in several bigger towns such as Gjilan, Ferizaj, Gjakova and Mitrovica, in order to offer students equal opportunity for studies closer to their home towns and regions where they are located. UP consists of a total of 13 faculties and all offer English Language Courses and it is presumed that the courses are under the umbrella of ESP. The general aim of the courses is primarily to enable students to practice and master the skills that they will be able to utilize in their further studies or future career and profession. In an ideal world the students should attain the needed skills and should be able to communicate effectively and fluently in the target situation, however this is not the case with the most of the UP students.

At the UP teachers face many challenges in teaching starting from materials, instruction tools, policies in place, class size, and proper methodology. This all reflects in the final outcome of the students' proficiency and motivation also.

3.2. Instruments

The respondents were presented and asked to answer the Albanian version of the questionnaire, as Albanian is

population's native language. It contained 62 questions and was designed to fit into four pages, and the time limit to answer it was calculated for 30 min total. Referring to foreign language research on questionnaires [1] the length of a questionnaire should not exceed 4 pages and it should to not exceed 30 min time limit for completing it. Following this recommendation the technical aspect of the questionnaire was closely considered, being an important part in designing the questionnaire and attracting respondents' interest and willingness to answer the questions. As Oppenheim [3] states the final design of the questionnaire has a significant role in attracting the respondents to answer it. Therefore the design also was created with special care: nested tables for the questions, separate part dedicated to different questions, multiple choice questions, yes /no questions, and finally the Likert type scale 1-5 (strongly disagree to strongly agree). The Likert type scale choice from 1-5 was selected due to the fact regional researchers have found it better for the region since the school grading system in the region follows the same pattern (1-5) 1- being unsatisfactory – 5 being excellent.

The questionnaire collected personal data from the respondents so that a test-retest would be conducted at a later stage in another extended longitudinal study. The personal data information was requested in the first page and it contained a total of 10 questions. The specific nature of ESP and particularly the context where the case study is conducted requires more data about the respondents. The demographic data (high school information, residency and background, other spoken languages, private language course attendance of the students) are vital in eliciting important information that cannot otherwise be retrieved only from the responses and from cross-checking. Thus, a full confidentiality and ethical aspects were considered and therefore were guaranteed and disclosed in the first page. The respondents were assured that the information will be used exclusively for the research purposes and it will not be disseminated by any means to any third party or to the public. Also the respondents were disclosed the information about the institution the researcher is affiliated with and where the study is taking place.

The original questionnaire was sent for review, editing and proofreading to professors, professional translators in the field of education and research in English as a foreign language. After several modifications the final draft was approved for translation. The translation into Albanian was done by the researcher herself, being a native Albanian speaker, and then it was sent for review to two other native speakers that are experts in editing translated documents in Albanian. The final draft after the last modification was printed in total 220 copies, and then distributed to the respondent as a hard copy. The primary goal of the distribution of the questionnaire was to test its validity and reliability before conducting the main study. For this purpose the University of Prishtina Law Faculty was selected, due to the reason that it is the faculty of primary interest for the study, and the researcher's own affiliation with the body granted easier access to its students.

4. Data analysis

In Table 1.1 a summary of the distribution of responses is provided. This information provides us with information on actual numbers and clearly shows that there is a skewed distribution towards strongly agree and agree for all respondents, except for question 6. This is in line with expectations and it proves that students are aware of basics of ESP courses and also of teacher's attitudes and their roles in an ESP setting.

Table 1: Distribution of Responses

	Responses					
Question	**1**	**2**	**3**	**4**	**5**	**Total**
Q1	1	4	31	78	92	206
Q2	0	2	24	63	116	205
Q3	3	4	40	50	107	204
Q4	2	14	50	63	80	209
Q5	5	8	34	84	67	198
Q6	48	30	37	45	47	207

Actual responses above are shown as relative frequencies (percentages) in table 2, and allows for a more direct comparison between the different rates of responses between different questions.

Table 2: Relative Frequencies

	Percentages					
Question	**R1**	**R2**	**R3**	**R4**	**R5**	**Total**
Q1	0.5%	1.9%	15.0%	37.9%	44.7%	100%
Q2	0.0%	1.0%	11.7%	30.7%	56.6%	100%
Q3	1.5%	2.0%	19.6%	24.5%	52.5%	100%
Q4	1.0%	6.7%	23.9%	30.1%	38.3%	100%
Q5	2.5%	4.0%	17.2%	42.4%	33.8%	100%
Q6	23.2%	14.5%	17.9%	21.7%	22.7%	100%

Moreover, this information is presented in Figure 1, where we are able to observe that on questions 1 to 4, majority of respondents answered strongly agree and agree [q1 44.7% strongly agree, 37.9% agree; q2 56.6% s.a., 30.7% a.; q3 52.5% s.a., 24.5% a.; q4 38.3% s.a., 30.1% a.] In question 5, this response rate is similar, but with more respondents answering agree (42.4% than strongly agree (33.8%). Question 6, however, has a more uniform distribution of responses, with largest number of respondents answering strongly disagree (23.2%), closely followed by strongly agree (22.7%), and agree (21.7%).

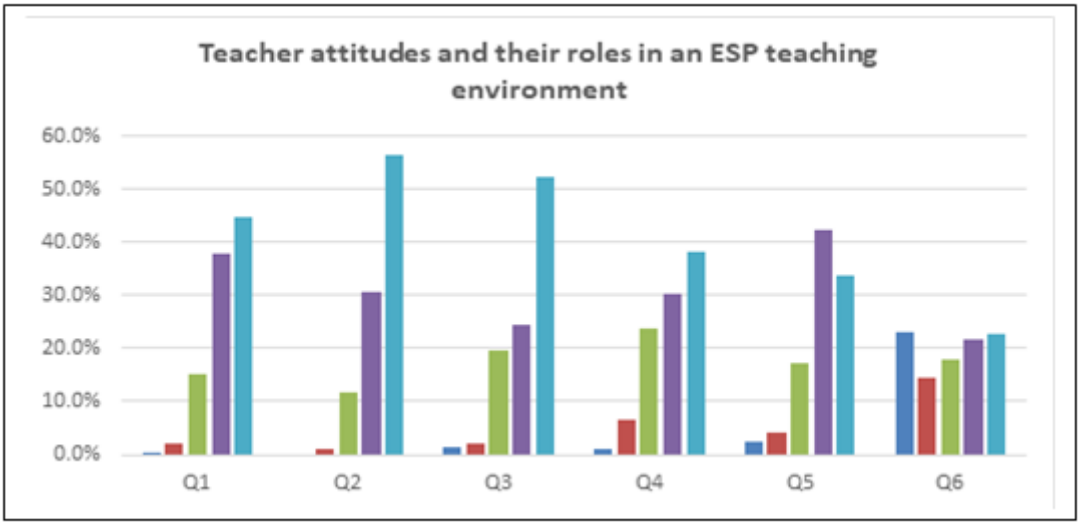

Figure 1: Percentages

Table 3, provides a summary of descriptive statistics for each question, that include median (middle value of response), mode (most frequent response), minimum (lowest) and maximum (highest) value of response, and both first and third quartiles. Quartiles are used to measure relative position, with first quartile indicating the response value that 25% of all respondents fall to the left (that is lower valued responses), and third quartile where 75% of all respondents fall to the left. In this case, if we look at question 6 responses, we see that 25% of all respondents answered 2 or less (or disagree or strongly disagree); and that 75% of all respondents answered 4 or less (or agree, neutral, disagree, or strongly disagree). Also, we use the values of first and third quartiles to measure the concentration of data. So (for question 6) we conclude that middle 50% of all respondents answered between disagree and agree.

Table 3: Descriptive Statistics

Descriptive Statistics						
Question	**Median**	**Mode**	**Min**	**Max**	**Quartile 1**	**Quartile 3**
Q1	4	5	1	5	4	5
Q2	5	5	2	5	4	5
Q3	5	5	1	5	4	5
Q4	4	5	1	5	3	5
Q5	4	4	1	5	4	5
Q6	3	1	1	5	2	4

Mann-Whitnney Test was used to test whether pairs of questions had the same median response (in other words if the data from the responses on different questions could have come from the same population.). Of particular interest in this section was the pairing of questions 2 (ESP teacher has excellent specialist knowledge) and 6

(The ESP course can be taught by any English Language teacher). The analysis of results produced a p-value of 0.000, which is less than the level of significance of 5% (0.05); indicating that median responses from the two questions were not equal. Mann-Whitney tests were carried out for all pairs of questions and are given in Table 4 below.

Table 4: Mann - Whitney U - Test: Paired Results

p-values of Analysis					
Question	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**
Q1	0.0149	0.5322	0.0106	0.0155	0.0000
Q2		0.1061	0.0000	0.0000	0.0000
Q3			0.0035	0.0045	0.0000
Q4				0.7946	0.0000
Q5					0.0000

Finally, Spearman rank correlation was used to measure the strength and the direction of the relationship between each pair of responses. Results of correlations are given in Table 5. For each pair of questions there are two values; top value provides the coefficient of correlation, and bottom value is the p-value that we measure against the level of significance (0.05) to denote if the paired correlations are significant or not – a p-value of less than 0.05 shows that the correlation is significant. Coefficient of correlation has a range of values from negative one to plus one (-1 to +1). A positive value indicates a direct relationship, whereas a negative value of the coefficient indicates an inverse relationship; with stronger relationships having values of closer to +1 or -1.

Table 5: Results of correlations

	P1.1	P1.2	P1.3	P1.4	P1.5
P1.2	0.3710				
	0.0000				
P1.3	0.3020	0.2150			
	0.0000	0.0000			
P1.4	0.2360	0.2540	0.1890		
	0.0000	0.0000	0.0020		
P1.5	0.1870	0.2750	0.2030	0.3210	
	0.0030	0.0000	0.0010	0.0000	
P1.6	0.0220	0.0380	-0.1050	0.1520	0.2140
	0.7170	0.5440	0.0880	0.0130	0.0010

Table 6: Spearman Correlation

Spearman	P4.6	P4.7	P4.8	P4.9	P4.10	P4.11	P4.12	P4.13	P4.14	P4.15
	P4.5.2	P4.6	P4.7	P4.8	P4.9	P4.10	P4.11	P4.12	P4.13	P4.14
P4.6	0.0000									
	1.0000									
P4.7	-0.6450	0.5000								
	0.2390	0.3910								
P4.8	-0.1670	0.3230	0.0000							
	0.7890	0.5960	1.0000							
P4.9	0.6450	0.5000	0.0000	0.3230						
	0.2390	0.3910	1.0000	0.5960						
P4.10	0.4080	0.7910	0.0000	0.6120	0.7910					
	0.4950	0.1110	1.0000	0.2720	0.1110					
P4.11	0.1480	0.8030	0.5740	0.2960	0.8030	0.7250				
	0.8120	0.1020	0.3120	0.6280	0.1020	0.1650				
P4.12	0.6450	0.5000	0.0000	0.3230	1.0000	0.7910	0.8030			
	0.2390	0.3910	1.0000	0.5960	*	0.1110	0.1020			
P4.13	0.0000	1.0000	0.5000	0.3230	0.5000	0.7910	0.8030	0.5000		
	1.0000	*	0.3910	0.5960	0.3910	0.1110	0.1020	0.3910		
P4.14	0.4080	0.7910	0.0000	0.6120	0.7910	1.0000	0.7250	0.7910	0.7910	
	0.4950	0.1110	1.0000	0.2720	0.1110	*	0.1650	0.1110	0.1110	
P4.15	0.4080	0.7910	0.0000	0.6120	0.7910	1.0000	0.7250	0.7910	0.7910	1.0000
	0.4950	0.1110	1.0000	0.2720	0.1110	*	0.1650	0.1110	0.1110	*

Cell Contents: Spearman rho
P-Value

5. Conclusion

The results of the assessment of the current questionnaire's validity and reliability proved positive, and it qualified as tool to be used in measuring the teacher attitudes and their roles in an ESP setting. Based on the previous studies most of the crucial parameters that make ease – to answer items were taken into consideration and the items proved to be respondent friendly. They were all answered in the given (short) time and in few (4) pages. This showed that the questionnaire can be successfully used for the main study and would give great results on the topic. The distribution of this questionnaire could elicit data that were not analyzed before in this context. It would bring answers to few questions and also raise some new questions for further research.

5.1. Limitations

There were several obstructions by teachers who were not interested in co-operation and not allowing their class

to be observed that at this point is considered as a limitation and therefore may have hindered the results of the pilot questionnaire. Due to this, alternative classes that were taught by other professors within the same faculties were visited, and the observations were successfully conducted. On the same note the mixed background, excessive number of freshmen students, and the early phase of piloting the instrument was another limitation of the study.

5.2. Recommendations

From the results of the pilot study the following recommendations are proposed for future practice and research:

1. The phase for piloting an instrument should be chosen carefully, as a very early start at the Universities doesn't include all the students, and those students that will be involved have very little knowledge about the course that they register to attend. Under the conditions students should be considered as a sample after the first semester, towards the end of the second semester. By this time they will be more familiar with the system in general and also they will be more experienced about the course and exams.

2. The sample size was too big for the purpose, instead smaller groups of selected students would give better results.

3. The piloting phase needs ample time and revisiting, as some of the items could have been rephrased or explained to the subjects.

References

[1]. Z. Dörnyei. Questionnaires in second language research: Construction, administration, and processing. Mahwah, New Jersey: Lawrence Erlbaum Associates, 2003.

[2]. J. D. Brown. Using surveys in language programs. Cambridge, U.K.: Cambridge University Press, 2001.

[3]. A. N. Oppenheim. Questionnaire design, interviewing, and attitude measurement. London: Pinter, 1992.

[4]. Z. Dornyei & T. Taguchi. Questionnaires in second language research: Construction, administration, and processing (2nd ed., Second Language Acquisition Research Series). Taylor & Francis, 2009.

[5]. J. A. Harkness. Survey methods in multinational, multiregional, and multicultural contexts. Hoboken, NJ: John Wiley & Sons, 2010.

[6]. R. W. Brislin, C. Cherrie & K. Cushner. Intercultural interactions: A pratical guide. London: Sage Publications, 1986.

[7]. J. D. Brown. Using surveys in language programs. Cambridge, U.K.: Cambridge University Press, 2001.

[8]. A. Mackey & S.M. Gass. Research methods in second language acquisition: A practical guide. Chichester, West Sussex: Wiley-Blackwell, 2012.